PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Basics of SVM and application to bearing fault diagnosis

by

Pavan K. Kankar
Mechanical Engineering Discipline

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

BASICS OF SVM

Linear Classification using SVM

- The goal of classification is to group items, that have similar feature values, into groups.

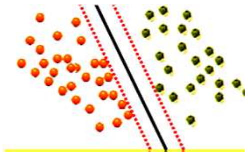


Figure: Separation between classes

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

BASICS OF SVM

- Fig. 1 shows rings and diamonds stand for these two classes of sample points, respectively; H is one of the separation planes.
- H1 and H2 are the planes those are parallel to H and, respectively, pass through the sample points closest to H in these two classes, the distance between H1 and H2 is defined as margin.

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

BASICS OF SVM

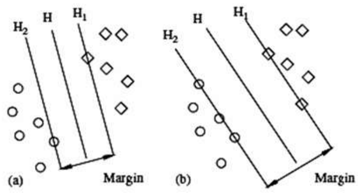


Figure: (a) A separation plane with small margin. (b) A separation plane with larger margin. A better generalization capability is expected from (b).

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

BASICS OF SVM

- Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression.
- A special property of SVMs is that they simultaneously minimize the generalization error and maximize the geometric margin.

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

BASICS OF SVM

Formulation

- Consider a training sample set {(x_i, y_i)}; i=1 to M, it is wished to determine, among all linear separation planes that separates input samples into two classes, which separation plane will have the smallest generalization error.

BASICS OF SVM

- Let us assume the samples can be classified into two classes namely positive class and negative class. Each of classes associate with labels be y_i = 1 for positive class and y_i = -1 for negative class, respectively.

BASICS OF SVM

- It is possible to determine the hyperplane *f(x)* = 0 that separates the given data

$$f(x) = w^T x + b = \sum_{i=1}^{M} w_i x_i + b = 0 \qquad (2)$$

- where *w* is M-dimensional vector and *b* is a scalar. The vector *w* defines a direction perpendicular to the hyperplane and the value of *b* moves the hyperplane parallel to itself, this value is sometimes called the bias (or threshold).

BASICS OF SVM

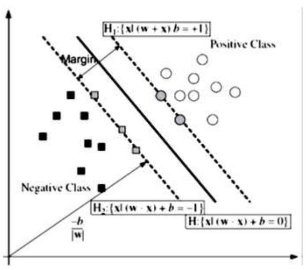


Figure: Hyperplane separating positive and negative class

BASICS OF SVM

- The decision function is made using sign f(x) to create separating hyperplane that classify input data in either positive class and negative class.

- A distinct separating hyperplane should satisfy the constraints

$$f(x_i) \ge 1 \quad if \quad y_i = 1,$$
$$f(x_i) \le -1 \quad if \quad y_i = -1$$

BASICS OF SVM

or it can be presented in complete equation as,

$$y_i f(x_i) = y_i(w^T x_i + b) \ge 1 \; for \; i = 1,2,\ldots,M \qquad (3)$$

BASICS OF SVM

- The SVM tries to place a linear boundary between the two different classes, and orientate it in such way that the margin is maximized.

- The boundary is placed in the middle of this margin between two planes. The nearest data points that used to define the margin are called ***support vectors***.

Contents

BASICS OF SVM

Margins of a Hyperplane

- The functional margin of an sample (xi,yi) with respect to a hyperplane (w,b) is the quantity

$$\gamma_i = y_i\left(\sum_{i=1}^{N} w_i x_i + b\right) \qquad (4)$$

$(y \in \{-1,+1\})$.

$\gamma_i > 0$ implies a correct classification of (xi, yi).

Contents

BASICS OF SVM

- The margin distribution of a hyperplane (w,b) w.r.t. a training set S is the distribution of the margins of samples in S.
- The geometric margin is the perpendicular Euclidean distance of the point to the hyperplane, i.e. $\gamma_i / \|w\|$

Contents

BASICS OF SVM

- The margin of a training set S is the maximum geometric margin over all hyperplanes.
- A hyperplane realizing this maximal is known as a maximal margin hyperplane.

Contents

BASICS OF SVM

- ***Theorem***: Vector *w* that minimizes $||w||^2$ under constraints

$$y_i f(x_i) = y_i(w^T x_i + b) \geq 1 \; for \; i = 1,2,\ldots,M \qquad (3)$$

is related to the vector that forms the optimal hyperplane by the equality.

Contents

BASICS OF SVM

Optimization Problem

- The optimal hyperplane separating the data can be obtained as a solution to the following optimization problem

$$\min imize \; \frac{1}{2} \| w \|^2 \qquad (5)$$

$$subject \;\; to \; y_i(w^T x_i + b) \geq 1, \quad i = 1,2,\ldots, M \qquad (6)$$

- The above is an optimization problem with a convex quadratic objective and only linear constraints. Its solution gives us the optimal margin classifier.

Contents

BASICS OF SVM

- To find out the solution of quadratic optimization problem saddle point of the Lagrange function has to be determine.
- The Langrangian function for optimization problem is given as

$$\text{minimize} \quad L(w,b,\alpha) = \frac{1}{2}\|w\|^2 - \sum \alpha_i [y_i(w^T x_i + b) - 1] \qquad (7)$$

where $\alpha_i \geq 0$ are the Lagrange multipliers.

Contents

BASICS OF SVM

- To find the saddle point one has to minimize this function over w and b and to maximize it over nonnegative Lagrange multipliers $\alpha_i \geq 0$. Setting the derivatives of L with respect to w and b to zero. We have:

$$\left.\begin{aligned} \frac{\partial L(w,b,\alpha)}{\partial w} &= w - \sum_{i=1}^{M} y_i \alpha_i x_i \\ \frac{\partial L(w,b,\alpha)}{\partial b} &= \sum_{i=1}^{M} y_i \alpha_i = 0 \end{aligned}\right\} \quad (8)$$

Contents

- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

BASICS OF SVM

- From conditions (8) it follows that the vector w that defines the optimal hyperplane, the equalities

$$\left.\begin{aligned} w &= \sum_{i=1}^{M} y_i \alpha_i x_i \\ \sum_{i=1}^{M} y_i \alpha_i &= 0 \end{aligned}\right\} \quad (9)$$

hold true. Substituting (9) into (7),

Contents

- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

BASICS OF SVM

$$W(\alpha) = \sum_{i=1}^{M} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{M} y_i y_j \alpha_i \alpha_j (x_i . x_j) \quad (10)$$

Notation L (w,b,α) is changed to W(α) to reflect the last transformation.

Contents

- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

BASICS OF SVM

To construct the optimal hyperplane, coefficients α_i are to be determine that maximize the function (10). Thus by solving the Dual optimization problem, the coefficients 'α_i^o' can be obtained which is required to express the 'w' to solve eq (5).

$$\max imize\ W(\alpha) = \sum_{i=1}^{M} \alpha - \frac{1}{2} \sum_{i,j=1}^{M} y_i y_j \alpha_i \alpha_j (x_i x_j) \quad (11)$$

$$\begin{cases} \alpha_i \geq 0, i = 1,2,..., M \\ \sum_{i=1}^{M} \alpha_i y_i = 0 \end{cases} \quad (12)$$

Contents

- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

BASICS OF SVM

SVM for the Non-Separable Case

Figure: Non-separable cases

Contents

- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

BASICS OF SVM

- To allow for the possibility of samples violating constraints, slack variables are introduced

$$\xi_i \geq 0, \quad i = 1,2,..., M \quad (13)$$

along with relaxed constraints

$$y_i (w^T x_i + b) \geq 1 - \xi_i, \quad i = 1,2,..., M \quad (14)$$

- A classier which generalizes well is then found by controlling both the classier capacity (via ||w||) and the number of training errors minimizing the objective function $\frac{1}{2} \| w \|^2 + C \sum_{i=1}^{M} \xi_i$ (15)

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

BASICS OF SVM

- The optimal hyperplane separating the data can be obtained as a solution to the following optimization problem

$$\text{min imize } \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{M}\xi_i \qquad (16)$$

$$\text{subject to} \begin{cases} y_i(w^T x_i + b) \geq 1 - \xi_i, & i = 1,2,..., M \\ \xi_i \geq 0, & i = 1,2,...M \end{cases} \qquad (17)$$

where C is a constant representing error penalty.

17-Sep-18 P.K. Kankar 33

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

BASICS OF SVM

- Rewriting the above optimization problem in terms of Lagrange multipliers, leads to the problem

$$\text{max imize } W(\alpha) = \sum_{i=1}^{M}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{M} y_i y_j \alpha_i \alpha_j (x_i x_j) \qquad (18)$$

$$\text{subject to} \begin{cases} 0 \leq \alpha_i \leq C, & i = 1,2,..., M \\ \sum_{i=1}^{M}\alpha_i y_i = 0 \end{cases} \qquad (19)$$

17-Sep-18 P.K. Kankar 34

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM in Bearing Fault Diagnosis

17-Sep-18 P.K. Kankar 35

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 1 [5]

- The key challenge to fault diagnosis is the improvement of diagnostics accuracy based on a given amount of information.
- This research investigates the possibilities of improving machinery diagnostics accuracy based on machine learning techniques.

17-Sep-18 P.K. Kankar 36

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 1

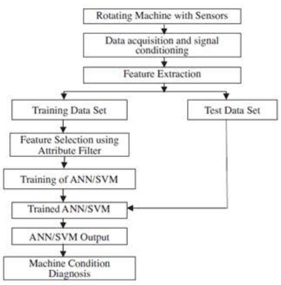


Figure: Flow chart of bearing health diagnosis

17-Sep-18 P.K. Kankar 37

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 1

- Following faults are introduced in the bearing:

(a) Outer race with crack and rough surface,

(b) Inner race with rough surface,

(c) Ball with corrosion pitting,

(d) Combination of above faults.

17-Sep-18 P.K. Kankar 38

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Application of ANN and SVM 1

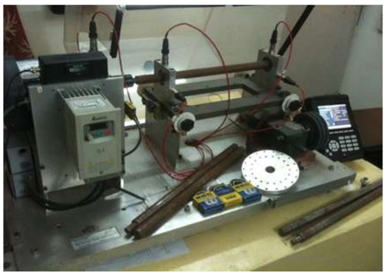
(a) (b)

Figure (a) Outer race with crack, (b) Ball with corrosion pitting

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Application of ANN and SVM 1



PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Application of ANN and SVM 1

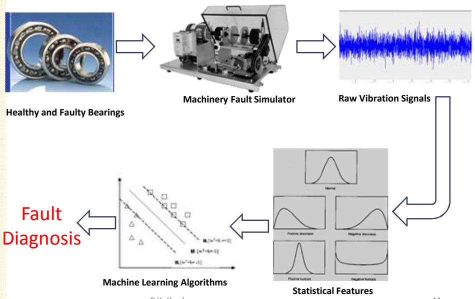


Figure: Procedure for Machine Condition Diagnosis

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Application of ANN and SVM 1

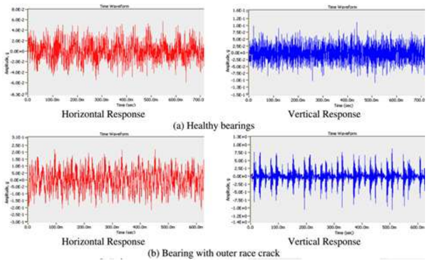


Figure: Vibration signals for various bearing conditions at 1000 rpm with no loader

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Application of ANN and SVM 1

Feature extraction and selection:

- A wide set of features are calculated from the vibration signals using statistics.
  - Range,
  - Mean value,
  - Standard deviation,
  - Skewness,
  - Kurtosis,
  - Crest factor.

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Application of ANN and SVM 1

ANN/SVM training and testing:

- Six statistical features are used each for horizontal and vertical response.
- Total 73 instances and 14 features are used for the study including statistical features, speed and number of loader used.

Contents

Application of ANN and SVM 1

Table 1 shows accuracy associated with each technique for fault classification. The correctly classified instances using test set for SVM and ANN are 73.9726% & 71.2329% respectively.

Table 1 Evaluation of the success of the numeric prediction

Parameters	SVM	ANN
Correctly Classified Instances	**54 (73.9726%)**	**52 (71.2329%)**
Incorrectly Classified Instances	19 (26.0274%)	21 (28.7671%)
Total Number of Instances	73	73

PDPM
Indian Institute of Information Technology,
Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 1

Summary

- This study presents a procedure for detection of bearing fault by classifying them using two machine learning methods, namely, ANNs and SVMs. Features are extracted from time-domain vibration signals using statistical techniques.
- The roles of different vibration signals, obtained with or without loader and at various speeds, are investigated.

PDPM
Indian Institute of Information Technology,
Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 1

Summary

- The results show the potential application of machine learning algorithm for developing a knowledge base system which can be useful for early diagnosis of defect for applying condition based maintenance to prevent catastrophic failure and reduce operating cost.

PDPM
Indian Institute of Information Technology,
Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 2 [6]

- The application of diagnostic techniques are to detect, and identifying a fault at the earliest possible stages of its initiation.
- This research investigates the possibilities of improving machinery diagnostics accuracy based on continuous wavelet transform (CWT) and machine learning techniques.

PDPM
Indian Institute of Information Technology,
Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 2

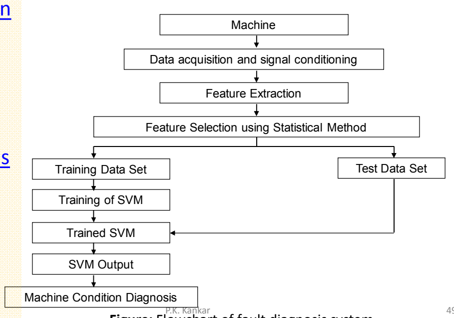


Figure: Flowchart of fault diagnosis system.

PDPM
Indian Institute of Information Technology,
Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 2

- Following five bearing conditions are considered for the study:
  1. Healthy bearings (HB)
  2. Bearing with spall on inner race (BSIR)
  3. Bearing with spall on outer race (BSOR)
  4. Bearing with spall on ball (BSB)
  5. Combined bearing component defects (CCD)

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Application of ANN and SVM 2

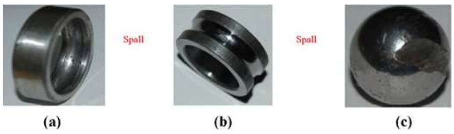


Figure: Bearing components with faults induced in them. (a) Outer race with spall, (b) inner race with spall, and (c) ball with spall

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Application of ANN and SVM 2

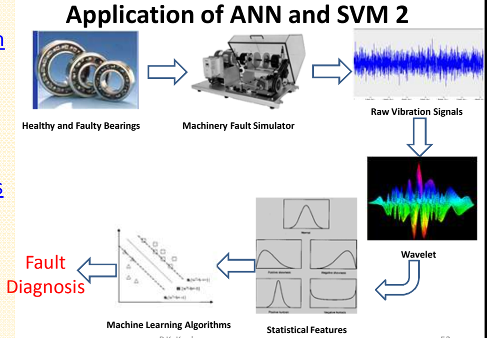


Figure 17 Procedure for Machine Condition Diagnosis

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Application of ANN and SVM 2

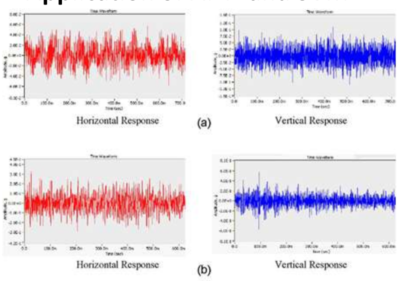


Figure: Vibration signatures for various bearing conditions at 1000 rpm. (a) Healthy bearings (HB), (b) bearing with spall on inner race (BSIR)

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Application of ANN and SVM 2

Feature extraction and selection:

- Three statistical features are calculated from the vibration signals using statistics.
  - Kurtosis,
  - Skewness,
  - Standard deviation.

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Application of ANN and SVM 2

- There are many wavelets those can be used with various machine learning methods for fault diagnosis.
- In order to extract the fault feature more effectively from the fault-induced transient vibration signals, an appropriate wavelet-base function should be selected.
- Maximum energy to Shannon entropy ratio criterion is used to select an appropriate wavelet and scale for feature extraction as shown in Fig 8.

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents
- Introduction
- Basics of ANN
- Basics of SVM
- Applications in Bearing Fault Diagnosis
- References

Figure 19 Flowchart for wavelet selection criterion

Raw Vibration Signals

3 real valued wavelets, 3 complex valued wavelets

128 scales, in 7th level

30 sample signals for bearing with spall in balls (BSB) | 30 sample signals for bearing with spall in inner race (BSIR) | 30 sample signals for bearings with combined defects (CCD) | 30 sample signals for healthy bearing (HB) | 30 sample signals for bearing with spall in outer race (BSOR)

Select one sample signal in a bearing condition | Select one Wavelet | One scale

Calculate Total Energy and Total Shannon Entropy of CWC for all scales for 30 samples in 5 bearing conditions

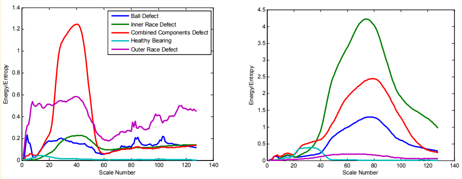
Calculate *Total Energy to Total Shannon Entropy Ratio* for each wavelet

Select wavelet which maximizes *Total Energy to Total Shannon Entropy Ratio* for Bearing Fault Diagnosis

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 2



(a) (b)

Figure: Plot between Energy to Shannon Entropy ratio vs. Scale number for Meyer wavelet Shaft running at (a)1000 RPM (b) 1500 RPM

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 2

- In the present study, training and testing of the classifiers such as Support Vector Machine (SVM) and Artificial Neural Network (ANN) have been carried out.
- Total 75 instances and 8 features are used for the study. These eight features are used as an input to train and test machine learning techniques.
- The signals are processed for analysis of machine condition diagnosis as shown in Fig. 10.

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 2

- Table 2 shows accuracy associated with each technique for fault classification. The correctly classified instances using test set for SVM and ANN are 98.6667% & 94.6667% respectively.

Table 2 Evaluation of the success of the numeric prediction

Parameters	SVM Test Set	ANN Test Set
Correctly Classified Instances	**74 (98.667%)**	**71 (94.667%)**
Incorrectly Classified Instances	1 (1.333%)	4 (5.333%)
Total Number of Instances	75	75

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

Application of ANN and SVM 2

Summary

- By using the proposed methodology, useful features can be extracted from the original data and dimensions of original data can be reduced by removing irrelevant features, so that the classifier can achieve a higher accuracy.
- The results show the potential application of proposed methodology with machine learning techniques for the development of on-line fault diagnosis system.

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

Contents

REFERENCES

[1] Chennakesava R. Alavala, “Fuzzy Logic and Neural Networks: Basic Concept and Applications”, New Age International (P) Limited, India.

[2] G. Dreyfus, “Neural Networks Methodology and Applications”, ISBN-10 3-540-22980-9 Springer Berlin Heidelberg New York.

[3] Micheli-Tzanakou, Evangelia, “Supervised and unsupervised pattern recognition: feature extraction and computational intelligence”, CRC Press, Boca Raton, Florida 33431.

[4] H.M. Ertunc, H. Ocak, C. Aliustaoglu, ANN- and ANFIS-based multi-staged decision algorithm for the detection and diagnosis of bearing faults, Neural Comput&Applic 22 (2013) (Suppl 1) S435–S446.

[5] P.K. Kankar, Satish C. Sharma, S.P. Harsha, “Fault diagnosis of ball bearings using machine learning methods”, Expert Systems with Applications 38 (2011) 1876–1886.

[6] P.K. Kankar, Satish C. Sharma, S.P. Harsha, “Fault diagnosis of ball bearings using continuous wavelet transform”, Applied Soft Computing 11 (2011) 2300–2312.

PDPM
Indian Institute of Information Technology, Design & Manufacturing Jabalpur

THANK YOU